USER GUIDE: HCUP SUPPLEMENTAL VARIABLES FOR REVISIT ANALYSES

Updated: February 16, 2015

Table of Contents

INTRODUCTION

The AHRQ Healthcare Cost and Utilization Project (HCUP) databases capture information on inpatient (IP), emergency department (ED) and ambulatory surgery (AS) encounters in U.S. community hospitals. These databases are often characterized as being "discharge-level" files, meaning that each record in a database represents one discharge abstract from a hospital setting, which can be an IP, ED, or AS visit. Thus, if the same individual visited the hospital multiple times in a given year, the HCUP databases would include separate records in the respective HCUP database for each visit. Many times researchers may be interested in knowing how many visits a distinct patient had rather than simply the number of overall hospital visits. Studying multiple visits is becoming increasingly common as hospital readmission rates are seen as an important indicator of the quality of medical care. To facilitate analyses that focus on multiple hospital stays by the same person, AHRQ created a set of supplemental variables that can be linked to the HCUP state-level databases to track multiple (repeat) patient visits in the hospital setting while adhering to strict privacy regulations. This user guide discusses the methodology used to develop these supplemental variables and how the information can be used with the HCUP databases. Appendices provide detail on which states, data types and years are available, in addition to verification statistics. Several SAS coding excerpts are also provided to facilitate the use of these files.

From data year 2003 to 2008, supplemental files called "HCUP Supplemental Files for Revisit Analyses" (herein referred to as the "Revisit Files") were created to include the revisit variables designed to augment the HCUP State databases. Beginning with 2009 data, the variables previously included in the Revisit Files were included in the Core file of the HCUP State Databases, when possible. The supplemental variables were labeled "revisit" rather than "readmission" intentionally. The term "revisits" implies multiple health care encounters for a particular patient that are not limited solely to IP stays. Most health care research on hospital *readmissions* has focused solely on the IP setting – i.e., tracking multiple hospital admissions in the IP setting by the same person. The HCUP revisit variables expand on traditional readmission analyses by allowing researchers to study multiple patient visits to the hospital, regardless of whether the visit resulted in a hospital readmission. In other words, these supplemental variables enable more than identification of hospital readmissions – they also enable tracking of patients admitted to the hospital following an ambulatory surgery (AS) or an emergency department (ED) visit and patients who made multiple trips to the ED. Note that revisits may occur for any reason (i.e., they may not be related) and can be separated by days or years. The determining factor in classifying health care events as revisits is that they represent services for the same individual.

In contrast, readmissions are sequential hospital admissions for a related reason, and usually within a specified time frame. Studying readmissions can be difficult as researchers must understand whether patients are admitted for expected follow-up treatment, or, conversely, for unexpected complications. In addition, multiple hospital visits for the same patient may, in fact, be unrelated - and, therefore, not considered a "readmission." Identifying readmissions requires specific criteria for the inclusion of events, such as type of condition and appropriate elapsed time. For example, a study of readmissions for congestive heart failure (CHF) may require that the principal reason for the hospitalization, ED visit, or ambulatory surgery is related to CHF and may also require that the time elapsed between events is not longer than predetermined number of days. The HCUP revisit variables contain key information, such as the days between multiple visits, that can assist analysts in making informed decisions about whether repeat patient visits qualify as readmissions.

The HCUP revisit variables, used in combination with HCUP state-level databases, enable analysts to link hospital visits that belong to a unique person, determine the elapsed time between visits, and evaluate valuable clinical information on the HCUP discharge abstract. These revisit data elements afford analysts the flexibility of performing patient-level analyses within and across hospital settings and time periods, without compromising patient confidentiality. Finally, these data elements allow the analyst to determine their definition of a readmission or revisit for the purposes of their study. This User Guide documents the creation of the revisit variables and provides guidance on how to best utilize them in revisit analyses.

HCUP Databases

HCUP develops and maintains a family of health care databases, related software tools, products and support services. HCUP features the largest collection of multi-year hospital care data in the United States, containing a wealth of all-payer, encounter-level information beginning in 1988. AHRQ relies on vital partnerships among Federal, State, and Industry associations to produce HCUP resources. HCUP databases integrate the data collected by state governments, hospital associations, private data organizations, and the Federal government to create a national health care information resource of IP, AS, and ED data.

The HCUP revisit variables are designed to be used exclusively with the HCUP state-level databases:

- The State Inpatient Databases (SID), which contain inpatient discharge records from non-Federal hospitals in participating states
- The State Emergency Department Databases (SEDD), which contain data from emergency department encounters from hospital-affiliated emergency departments
- The State Ambulatory Surgery and Services Databases (SASD), which include data from ambulatory care encounters from hospital-owned and sometimes non-hospital-owned ambulatory surgery sites.

The revisit variables are unique within state and data year. Prior to 2009 data, users will need to merge the data elements on the HCUP Revisit Files to the corresponding SID, SASD, or SEDD for any analysis (further described in the section on Using HCUP Revisit Variables). Starting with the 2009 data, the revisit variables are on the HCUP Core file.

Note: HCUP revisit variables are available for some, but not all, state-level databases – SID, SASD, and SEDD – starting in calendar year 2003. Appendix A lists the availability of the revisit variables by state, year, and data type.

DEVELOPMENT OF THE HCUP REVISIT VARIABLES

Development of the HCUP revisit variables required state-level HCUP files to contain a unique encrypted patient identifier which enable tracking of unique patients within and across years. Only some of the HCUP statewide data organizations provide this information to HCUP. Each state employs a distinct methodology in producing their encrypted person numbers.

Verified Patient Identifiers (visitLink)

An instrumental part of constructing the HCUP revisit variables was verifying that the encrypted person numbers accurately represented a unique person in the HCUP state databases. As part of the verification process, the patient's date of birth and sex were used to qualify the encrypted patient numbers and uniquely identify a person. A new verified person number (**visitLink**) was assigned for each unique combination of the qualifying information (encrypted person number, date of birth, and sex). Consider the following example: Five records have the same encrypted person numbers, but two records have one date of birth and sex, and the remaining three records have a different, but consistent, date of birth and sex. The two records with identical identifying information have one value of **visitLink,** and the other three records have a different value of **visitLink**. Appendix B contains examples of the assignment of **visitLink** for different scenarios.

No verified person number is assigned if any of the three pieces of information was missing (i.e., **visitLink** is missing). Additionally, no verified person number is assigned if there were more than 40 hospital visits in a given calendar year with the same qualifying information. This second qualification excluded less than 0.5 percent of the person numbers and aimed to eliminate person numbers used for multiple people. Appendix B contains examples of the assignment of **visitLink** for different scenarios.

While the term "verified person number" is used to describe the information in the HCUP data element **visitLink**, the values are not recognizable as specific patient information. **VisitLink** does not include the values of the encrypted person number, date of birth, or sex.

The **visitLink** variable is created each year for all discharges across all available databases for the particular state. For example, if a state provides SID, SASD, as well as SEDD, the **visitLink** is verified for all discharges across all three databases for that data year and previous years of data, as appropriate.

Calculating the Days to Event (daysToEvent)

For a verified person number (i.e., non-missing **visitLink**) with more than one hospital visit, the elapsed days between visits are calculated as the difference between the two visit dates. This information is often useful for determining readmissions for a specific condition (i.e., 30-day readmissions, 7-day readmissions, etc.) While this information is critical for defining readmissions, the use of admission and discharge dates is highly restricted per Health Insurance Portability and Accountability Act (HIPAA) guidelines.

To comply with HIPAA guidelines and ensure patient confidentiality, no "date" information is released on the HCUP revisit variables. A timing variable (**daysToEvent)** was calculated consistently for each verified person number (**visitLink**) based on a randomly assigned "start date." Each verified person number is assigned a unique start date that is used to calculate **daysToEvent** for all visits associated with that **visitLink** value. The variable **daysToEvent** is the difference between the visit's admission date and the start date associated with the **visitLink**.

The calculation of days between visits is the difference of **daysToEvent** between two selected visits for a unique verified person number (**visitLink**). For example, consider a patient with congestive heart failure that has a hospital admission on 1/10/2008 and an ED visit on 1/25/2008. If the **daysToEvent** value is "9" for the 1/10/2008 admission and the **daysToEvent**

value is “24” for the 1/25/2008 ED visit, then the number of days between the start of the first visit and the start of the second visit is 15 days (24 – 9 = 15). It should be noted that often readmission analyses consider the time between the end of one admission and the start on the next admission. To adjust for the length of the admission, subtract the length of stay from the difference. In the example, above, it the first admission had a length of stay of 2 days then the number of days between the end of the first visit and the start of the second visit is 13 days (24 – 9 – 2 = 13).

The lowest value of **daysToEvent** will be on the first or earliest event for a patient. It is important to remember that if patient A has a value of 605 for **daysToEvent** and patient B has a value of 300 for **daysToEvent**, patient B’s event did not necessarily take place prior to patient A’s event – in fact, Patient B's **daysToEvent** value has no relation to Patient A's **daysToEvent** value. Because of the use of a random start date in the calculation of **daysToEvent**, the value of **daysToEvent** cannot be compared across patients. Appendix B contains examples of the assignment of **daysToEvent** for different scenarios.

HCUP REVISIT VARIABLES

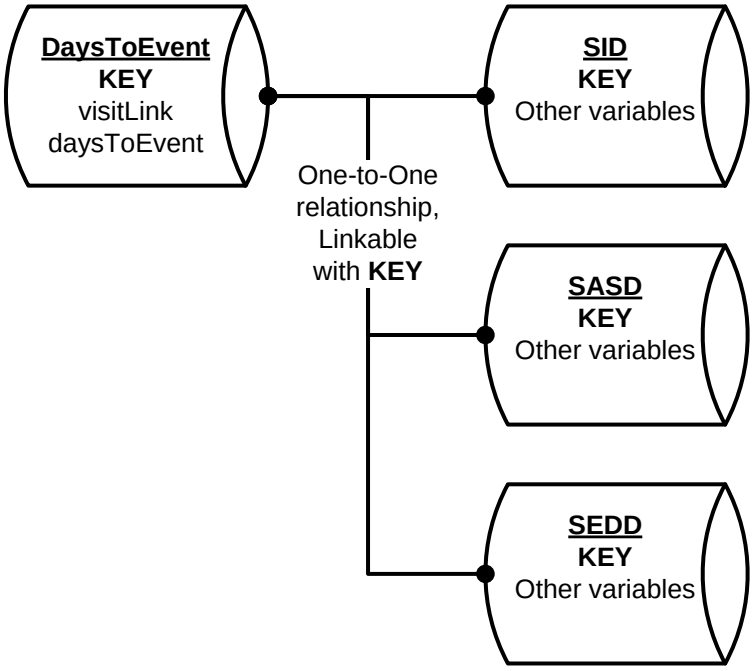
The HCUP revisit variables include only two data elements:

- **visitLink**: linkage variable for all events associated with a unique patient that is assigned during construction of the supplemental revisit variables and based on a unique combination of person numbers, date of birth, and sex
- **daysToEvent**: the number of days from a randomly chosen "start date" to the admission date for a specific health care visit for an individual. The start date is randomly assigned for each unique patient. As a result, **daysToEvent** will be consistently calculated for all of a patient’s linked events, regardless of year (i.e., all visits with the same value of **visitLink**). The daysToEvent variable is assigned during construction of the supplemental revisit variables.

From 2003-2008 the HCUP revisit variables are stored in separate state- and year-specific files (called Revisit Files) that can be linked to the corresponding state’s HCUP state data (SID, SASD, and SEDD) for that year (Figure 1). For example, for Nebraska there is one HCUP revisit linkable file for the 2006 data year. Researchers can add the **visitLink** and **daysToEvent** data elements to the 2006 Nebraska SID, SASD, and/or SEDD files by linking on the **KEY** data element.

Starting in data year 2009, the revisit variables (**visitLink** and **daysToEvent)** are stored in the Core file of the SID, SASD, and/or SEDD, so there are no separate Revisit Files.

Figure 1. Relationship Between HCUP Databases



AVAILABILITY OF VERIFIED PATIENT IDENTIFIERS VARIES BY STATE

The availability of verified person numbers for specific populations and settings varies by state and should be considered prior to any analysis. Table 1 shows the range of the percentage of verified revisit information across 15 states in 2005-2006 for selected patient characteristics, expected payer, and hospital characteristics.

Table 1. Range of Percentage of Records with Verified Revisit Information

		All Events			
		Min	**Q1**	**Median**	**Max**
Overall		65.4%	84.7%	90.5%	100.0%
By Patient Characteristics					
Age Group	0	11.0%	26.8%	49.1%	100.0%
	1-17	35.1%	54.1%	73.3%	100.0%
	18-44	69.5%	92.6%	95.9%	100.0%
	45-64	69.2%	96.4%	97.5%	100.0%
	65+	69.8%	97.2%	98.8%	100.0%
Sex	Male	64.3%	82.6%	88.9%	100.0%
	Female	66.2%	86.6%	91.8%	100.0%
Patient Income	Quartile 1 (lowest)	59.3%	87.2%	91.5%	100.0%
	Quartile 2	66.9%	86.7%	91.5%	100.0%
	Quartile 3	70.1%	86.1%	90.9%	100.0%
	Quartile 4 (highest)	70.5%	83.4%	88.9%	100.0%
Expected Payer	Medicare	69.8%	97.2%	99.1%	100.0%
	Medicaid	55.4%	71.7%	85.9%	100.0%
	Private insurance	66.5%	82.2%	89.4%	100.0%

		All Events			
		Min	**Q1**	**Median**	**Max**
	Self-pay	66.9%	79.8%	88.2%	100.0%
	No Charge	54.2%	79.0%	86.5%	100.0%
	Other	44.0%	87.6%	93.7%	100.0%
By Hospital Characteristics					
Hospital Ownership	Government, nonfederal	72.5%	81.7%	90.5%	100.0%
	Private, not-profit	67.9%	84.7%	90.0%	100.0%
	Private, invest-own	57.5%	86.0%	92.5%	100.0%
Hospital Location	Large central metropolitan	78.0%	82.4%	87.8%	100.0%
	Large fringe metropolitan	36.3%	83.1%	89.2%	100.0%
	Medium metropolitan	74.3%	85.4%	92.0%	100.0%
	Small metropolitan	45.6%	87.8%	94.4%	100.0%
	Micropolitan	69.9%	88.7%	95.4%	100.0%
	Noncore	35.1%	83.4%	94.8%	100.0%
Hospital Bed Size	<100	56.7%	87.1%	92.5%	100.0%
	100-299	68.7%	84.3%	91.1%	100.0%
	300-499	70.1%	89.7%	92.2%	100.0%
	500+	72.7%	84.3%	92.5%	100.0%

Source: HCUP State Inpatient Databases, 15 States, 2005-2006

In most cases, verification rates across patient and hospital characteristics and across selected diagnosis and procedure categories were consistent with the overall verification rates. For example, the first and second quartiles of the verification percentage overall were 84.7 percent and 90.5 percent, respectively. The first and second quartiles of the verification percentage for patients from hospitals in large fringe metropolitan areas were 83.1 percent and 89.2 percent, respectively.

Some notable exceptions include:

- Newborns (age 0) – The median of the verification rates across the 15 states was only 49.1 percent.
- Children and adolescents (age 1-17) – The first quartile for verification rates was 54.1 percent and the median was 73.3 percent. A separate analysis examined whether verification rates were better for certain ranges of children, such as adolescents or teens. There was no specific range of pediatric ages between 1 and 17 that were markedly better in terms of the percentage verified person numbers.
- Expected payer of Medicaid and No Charge – The first quartile for both was less than 80 percent and the median was about 86 percent.

Revisit/readmission analyses for pediatric conditions and certain payers may only be appropriate in selected states.

SELECTING STATES FOR A REVISIT/READMISSION ANALYSIS

When selecting which states to use for a revisit analysis, please reference the following resources:

- Appendix A provides the list of all states, years, and databases with HCUP revisit variables.
- Appendix C lists states that have inconsistent coding across data years of the source person numbers provided by the HCUP Partners and should not be used for analyses that span certain years.
- Appendix D provides verification rates by state and year that should be used to determine which HCUP states are best for specific types of revisit or readmission analysis.

USING THE HCUP REVISIT VARIABLES

Using the HCUP revisit variables involves four basic steps

1. For a given state and year, merge the HCUP Revisit File with the corresponding SID, SASD, or SEDD by the data element **KEY** to add the revisit data elements **visitLink** and **daysToEvent**. This step is only needed for data years 2003-2008. Beginning in data year 2009, the data elements **visitLink** and **daysToEvent** are included on the Core file of the SID, SASD, and SEDD, when possible.
2. Select patients of interest.
3. Use **visitLink** to identify all events for a patient. The same unique value of **visitLink** is coded on all records for an individual patient. Records with missing values for the **visitLink** variable will be a mixture of patients with unknown revisit information. It may be appropriate to exclude these records from the analysis.
4. Use **daysToEvent** to sequentially order the visits for a patient and to calculate the time between two visits for a patient. If the **daysToEvent** is 5 on one event and 35 on another, the time between the *start* of each event is 30 (35-5) because **daysToEvent** is based on the admission date. If you want to consider the time between the end of the first event and the start of the second event, the length of stay for the first event needs to be subtracted. If the length of stay on the first event is two, then the number of days between is 28 (35-5-2 = 28)

Usage Examples

Use of the HCUP revisit variables is relatively straightforward. Below are three examples of applying these variables to research topics.

Usage Example #1: Assigning Patient Characteristics

Researchers may want to group patients by specific patient characteristics, such as a patient's age or insurance status. When a patient's health care experience includes more than one hospital event, categorizing the patient may be problematic. This difficulty arises because some patient characteristics may change over time. To assign attributes based on when a person began receiving services, consider the following steps:

- Data should first be grouped by patient, in service date order (specifically, the HCUP state-level data file, merged with the HCUP Revisit File if prior to 2009, and then sorted by **visitLink** and **daysToEvent**).

- All records for a patient are then sequentially examined in order to select the first valid, non-missing value for each patient characteristic (age, sex, race, income quartile, location, expected payer).
- The selected attribute(s) are then applied to all events for the patient.

A SAS coding example of how attributes can be assigned is shown in Appendix E.

Usage Example #2: Revisits for Selected Patients

This example counts the number of related events for selected patients with a specific diagnosis and calculates a number of statistics, including days between the initial event and the first subsequent event, by setting. This example focuses on revisits for diabetes, but can easily be adapted to any diagnoses.

The example looks for a "clean period," measured in months, with no hospital events for an individual patient for the specified condition. Use of a "clean period" for counting readmissions is optional. Sometimes when identifying an episode of care, rather than straight utilization, a period of time during which the patient has not been admitted or treated is required. The first event after the "clean period" is considered the index event. Any event in a predetermined period of time after the index event becomes part of the "episode."

For illustrative purposes, we selected "diabetes mellitus with complications" (CCS diagnosis category 50) as the condition and required a clean period of 6 months. The steps are:

1. Combine the event and revisit data
    a. Limit data to linkable patients (a non-missing **visitLink** available)
    b. Select all events with the specified condition (diabetes)
2. Sort the combined events into patient (**visitLink**) and service sequence (**daysToEvent**) order
3. Find individuals with two or more events
    a. Find patients with a clean period before their first diabetes event
        i. On the first event for a patient, the service must be after clean period, defined as the first six months of the data year
        ii. If the first event was prior to the clean period month, look for a clean period on subsequent events by testing the number of days between the current event and the preceding event
    b. When a clean period is identified
        i. Count the number of events after the clean period
        ii. Determine the settings of the first and second events, and calculate the number of days between the first and second event
    c. Summarize the processing counts
4. Calculate statistics (distribution) for the number of patient events
5. Summarize revisits by the initial and second service settings.

SAS programming code for this example is found in Appendix E.

Usage Example #3: Preceding Visits to any Hospital Setting for Selected Patients

This example identifies patients' hospital events that precede CABG surgery (CCS procedure category 44), regardless of service setting, and summarizes counts by principal diagnosis. The steps are:

1. Combine the event and revisit data
    a. Limit data to linkable patients (a non-missing **visitLink** available)
    b. Identify events with the specified procedure (CABG)
2. Sort the combined events into patient (**visitLink**) and service sequence (**daysToEvent**) order
3. For patients who received CABG surgery, select all events prior to the surgery
4. Summarize prior events by primary diagnosis and setting

SAS programming code for this example is found in Appendix E.

Cautionary Note: Transfers and Possible Duplicates

The HCUP revisit variables allow an analyst to identify which records in the SID, SASD, and SEDD belong to the same person, as well as the time between events for that person. An analyst still must decide how to handle the following two types of scenarios:

- Transfers – when a patient is transferred from one acute care hospital to another
- Duplicates – when a record for the same event occurs twice in the HCUP file.

In the SID, there will be two different records if a patient is transferred from one hospital to another. The following can be used to identify the two SID records:

- Same person (**visitLink** is the same on two records)
- Disposition indicating transferred out (**DISPuniform** = 2)
- Admission source indicating transfer in (**ASOURCE** = 2)
- Discharge date of one record is the same as the admission date of another (**daysToEvent** plus the length of stay of the first record equals the daysToEvent of the second record)
- Different hospital (**DSHOSPID** is different).

Analysts conducting patient-level analyses need to decide how best to use the above information to identify transfers. The coding of admission source and discharge disposition is not always consistent with the timing of events identified by **daysToEvent** (i.e., **daysToEvent** may identify two records as two parts of a transfer, but either the disposition or admission source is not coded as such). Table 2 demonstrates the range in the percentage of discharges identified as transfers using different schemes.

Table 2. Range of Percentage of Inpatient Discharges Identified as Transfers

Scheme to identify transfers	Minimum Value Across 15 States	Maximum Value Across 15 States
Percentage of records identified as transfers using one source of information:		
Dates	1.80%	6.41%
Discharged as a transfer to another acute care hospital (DISPuniform=2)	1.36%	3.11%
Admitted as a transfer from an acute care hospital (ASOURCE=2)	0.85%	5.30%
Percentage of records identified as transfers using two sources of information:		
Dates and DISPuniform=2	0.70%	2.07%
Percentage of records identified as transfers using all three sources of information:		
Dates, DISPuniform=2, ASOURCE=2	0.22%	1.50%

Source: HCUP State Inpatient Databases, 15 States, 2006

For some analyses it may be best to combine the two records from a transfer into one by summing the lengths of stay and total charges and combined diagnoses and procedures.

The HCUP SID, SASD, and SEDD occasionally have multiple records for the same person (**visitLink**) with the same **daysToEvent** and length of stay (**LOS**). These duplicate records may or may not have the similar charge and diagnostic information. Analysts should decide how best to handle such records.

In addition, HCUP made an explicit decision to duplicate records across the SEDD and SASD when a record indicated that the patient received services in both settings. In this case, the analyst will need to decide how to include and account for these cases. The effect of these duplicated records varies by state from less than 1% in California SASD to about 15% in Tennessee SASD.

APPENDIX A: AVAILABLE HCUP SUPPLEMENTAL VARIABLES FOR REVISIT ANALYSES

HCUP revisit variables are available for the following states, years, and databases. For data years 2003-2008, the supplemental Revisit Files must be linked to corresponding HCUP SID, SASD, or SEDD for any analysis. Starting with 2009 data, the revisit variables are included in the Core file, when possible.

Information on some HCUP databases are to be determined (TBD) after HCUP data processing.

State and Data Type	Data Year										
	2003	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Arizona											
SID	Yes	Yes	Yes	Yes	Yes	--	--	--	--	--	--
SEDD	--	--	Yes	Yes	Yes	--	--	--	--	--	--
Arkansas											
SID	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD
California											
SID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD	TBD
SASD	--	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD	TBD
SEDD	--	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD	TBD
Florida											
SID	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SASD	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SEDD	--	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Iowa											
SID	--	--	--	--	--	--	Yes	Yes	Yes	Yes	Yes
SASD	--	--	--	--	--	--	--	Yes	Yes	Yes	Yes
SEDD	--	--	--	--	--	--	--	Yes	Yes	Yes	Yes
Maryland											
SID	--	--	--	--	--	--	--	--	--	Yes	TBD
SASD	--	--	--	--	--	--	--	--	--	--	TBD
SEDD	--	--	--	--	--	--	--	--	--	--	TBD
Massachusetts											
SID	--	--	--	--	--	--	--	Yes	Yes	Yes	TBD
SEDD	--	--	--	--	--	--	--	Yes	Yes	Yes	TBD
Mississippi											
SID	--	--	--	--	--	--	--	Yes	Yes	--	--
Nebraska											
SID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SASD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SEDD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

State and Data Type	Data Year										
	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Nevada											
SID	Yes	Yes	Yes	Yes	Yes	--	--	--	--	--	TBD
New Mexico											
SID	--	--	--	--	--	--	Yes	Yes	Yes	Yes	TBD
New York											
SID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD
SASD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD
SEDD	--	--	--	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD
North Carolina											
SID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	--	--	--
SASD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	--	--	--
SEDD	--	--	--	--	Yes	Yes	Yes	Yes	--	--	--
Utah											
SID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD	TBD
SASD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD	TBD
SEDD	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD	TBD
Vermont											
SID	--	--	--	--	--	--	--	--	Yes	Yes	Yes
SASD	--	--	--	--	--	--	--	--	Yes	Yes	Yes
SEDD	--	--	--	--	--	--	--	--	Yes	Yes	Yes
Washington											
SID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	TBD
Wisconsin											
SID	--	--	--	--	--	--	--	--	--	--	Yes
SASD	--	--	--	--	--	--	--	--	--	--	Yes
SEDD	--	--	--	--	--	--	--	--	--	--	Yes

APPENDIX B: EXAMPLE OF ASSIGNMENT OF VISITLINK AND DAYSTOEVENT

The following table lists examples of the assignment of **visitLink** and **daysToEvent** in different scenarios.

Observation	Example PNUM	Example DOB	SEX	Example visitLink	Calculation for daysToEvent	Notes
1	A	15-Jan-1960	M	11111	Calculated using same randomly selected start date for observations 1 and 2	Same PNUM, DOB, and sex for observations 1 and 2, therefore same visitLink
2	A	15-Jan-1960	M	11111		
3	A	15-Jan-1960	F	11112	Calculated using different randomly selected start date	Same PNUM and DOB as observations 1 and 2, but different sex, therefore different visitLink
4	B	1-May-1940	F	11113	Calculated using same randomly selected start date for observations 4 through 6	Same PNUM, DOB, sex for observations 4 through 6, therefore same visitLink
5	B	1-May-1940	F	11113		
6	B	1-May-1940	F	11113		
7	B	15-Jun-1945	F	11114	Calculated using same randomly selected start date for observations 7 through 9	Same PNUM and sex as observations 4 through 6, but different DOB, therefore different visitLink
8	B	15-Jun-1945	F	11114		
9	B	15-Jun-1945	F	11114		
10	C	1-Dec-1980	M	11115	Calculated using same randomly selected start date for observations 10 and 11	Same PNUM, DOB, sex for observations 10 and 11, therefore same visitLink
11	C	1-Dec-1980	M	11115		
12	C	Not Available	M	Missing	Missing	Same PNUM and sex as observations 10 and 11, but missing date of birth, therefore visitLink and DaysToEvent are missing

APPENDIX C: CONSISTENCY OF PERSON IDENTIFIERS ACROSS YEARS

The HCUP data element **visitLink** is derived from encrypted person numbers provided by the HCUP Partner. Partners sometimes change their coding scheme between data years, which in turn causes a discontinuity in **visitLink.** The table below lists the percentage of unique values of **visitLink** that appear in consecutive data years of the SID. If the percentage is low or different than other pairs of years, it is a good indication that the **visitLink** cannot be used to track patients across those data years. A dash indicates that **visitLink** is not available in one or both years.

To better understand how to interpret the table below, consider the following example. In Washington, 16 percent of the **visitLink** values in 2003 also appeared in 2004. This is a good indication that **visitLink** can be used to track WA patients between 2003 and 2004. The percentage of overlap in **visitLink** is 17 percent in 2004-2005 and 2005-2006 for WA. This is a good indication that **visitLink** can be used to track WA patients from 2003 through 2006. In contrast, between the next two data years, 2006 and 2007, the percentage of overlap in **visitLink** is zero for WA. This indicates that **visitLink** should *not* be used to track WA patients from 2006 into 2007, from 2007 into 2008, etc. In addition, the percentage of overlap is 100 percent between 2011 and 2012 for WA. The exact same values of visitLink were used for different people in these two data years; do not use the two years of WA SID (2011 and 2012) together.

Information on some HCUP databases are to be determined (TBD) after HCUP data processing.

	Percentage of visitLink Values Reported in Consecutive Data Years of the State Inpatient Databases (SID)									
State	**2003-2004**	**2004-2005**	**2005-2006**	**2006-2007**	**2007-2008**	**2008-2009**	**2009-2010**	**2010-2011**	**2011-2012**	**2012-2013**
AR	--	22%	22%	22%	22%	23%	23%	23%	23%	24%
AZ	12%	12%	11%	13%	--	--	--	--	--	--
CA	22%	22%	22%	22%	22%	22%	22%	23%	23%	23%
FL	--	0%	24%	24%	24%	25%	25%	26%	25%	26%
IA	--	--	--	--	--	--	20%	20%	20%	0%
MA	24%	25%	25%	25%	25%	25%	26%	26%	26%	TBD
MD	--	--	--	--	--	--	--	--	--	8%
MS	--	--	--	--	--	--	--	25%	--	--
NC	20%	21%	20%	21%	21%	21%	22%	--	--	--
NE	18%	16%	17%	18%	18%	18%	18%	18%	18%	17%
NM	--	--	--	--	--	--	4%	2%	0%	0%
NV	16%	16%	17%	18%	--	--	--	--	--	
NY	0%	0%	20%	21%	20%	20%	20%	20%	20%	20%
UT	14%	14%	14%	14%	14%	14%	14%	14%	14%	TBD
VT	--	--	--	--	--	--	--	--	18%	19%
WA	16%	17%	17%	0%	0%	0%	0%	0%	100%	0%

APPENDIX D: CONSISTENCY OF VERIFIED REVISIT INFORMATION

The consistency of the verified person numbers is evaluated when the HCUP revisit variables are created for a state. The year-specific tables on the following pages detail the number of total records in the SID, SEDD, or SASD and the percentage of records with a verified person number (**visitLink**).

The tables allow the analyst to determine the best possible states for a revisit analysis. Researchers should use the HCUP revisit variables with caution when looking at revisits for specific patient populations that have a low percentage of verified person numbers. If studying pediatric conditions, consider states with a high percentage of verified person identifiers for ages under 18. If a proposed study is specific to other patient or hospital characteristics, generate statistics on the percent verified by the study focus and select states with a high percentage of verified person identifiers.

	State Inpatient Databases, 2013							
	Age in Years							
	0		1-17		18-64		65+	
State	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
FL	239,661	5.3	109,438	63.7	1,279,462	95.4	1,044,819	98.6
IA	41,761	100.0	9,602	100.0	136,085	100.0	130,447	100.0
NE	28,016	100.0	8,464	100.0	97,172	100.0	71,068	100.0
VT	6,109	10.5	1,240	76.9	23,771	94.5	20,090	95.7
WI	71,294	100.0	26,690	100.0	295,110	100.0	218,637	100.0

	State Emergency Department Databases, 2013							
	Age in Years							
	0		1-17		18-64		65+	
State	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
FL	232,784	39.8	1,593,749	61.4	4,738,985	94.6	1,009,708	97.4
IA	34,918	100.0	200,448	100.0	617,608	98.8	183,239	99.9
NE	17,217	100.0	99,025	100.0	286,720	99.4	77,780	99.9
VT	3,848	27.3	39,481	64.3	157,738	91.6	41,558	92.6
WI	50,098	100.0	302,630	100.0	1,071,720	99.3	279,722	99.9

State	State Ambulatory Surgery and Services Databases, 2013							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
FL	10,418	47.8	132,402	60.3	1,445,382	92.3	1,311,089	95.1
IA	3,702	100.0	32,328	100.0	267,137	99.9	157,383	99.9
NE	1,982	100.0	16,793	100.0	104,911	99.9	56,198	99.9
VT	7,324	29.1	78,756	69.7	915,232	92.9	617,105	92.2
WI	5,797	99.0	66,228	98.4	712,191	96.9	390,419	97.4

State	State Inpatient Databases, 2012							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	41,844	43.7	20,706	85.5	191,447	97.4	152,240	98.4
FL	239,281	5.6	109,906	66.5	1,279,144	95.8	1,042,061	98.7
IA	41,766	5.3	10,423	38.1	141,470	74.8	134,633	83.6
MA	81,881	2.9	30,138	43.6	400,621	92.7	306,624	97.3
MD	74,727	11.7	21,324	31.7	368,940	19.8	230,022	17.5
NE	28,087	100.0	9,948	100.0	96,568	100.0	71,021	100.0
NM	27,572	97.2	11,005	97.8	100,258	97.6	60,143	98.8
NY	265,864	96.0	106,974	97.8	1,295,664	96.5	860,854	99.6
VT	6,081	10.0	1,305	78.0	24,327	95.9	20,467	96.7
WA	87,051	100.0	24,194	100.0	318,154	100.0	211,472	100.0

State	State Emergency Department Databases, 2012							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
FL	229,459	42.3	1,567,860	64.0	4,631,580	95.1	956,475	97.6
IA	32,095	22.4	216,686	40.6	631,899	81.9	180,576	88.7
MA	50,426	17.9	442,934	44.5	1,740,138	88.4	316,134	92.8
NE	19,667	100.0	107,154	100.0	275,176	99.0	73,366	100.0
NY	195,105	86.7	1,358,648	87.8	4,504,692	95.4	756,454	98.8
VT	4,110	30.7	41,541	68.9	160,923	93.3	41,329	95.5

State	State Ambulatory Surgery and Services Databases, 2012							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
FL	10,451	48.9	131,864	63.2	1,460,695	94.2	1,292,648	96.4
IA	3,405	15.4	32,343	37.3	263,562	74.1	152,964	79.7
NE	2,375	100.0	18,800	100.0	102,210	99.9	53,539	100.0
NY	258,723	83.7	1,260,648	83.7	7,917,091	85.9	2,846,115	94.9
VT	7,534	27.0	85,133	72.4	918,812	94.2	587,381	93.8

State	State Inpatient Databases, 2011							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	42,997	48.1	20,727	85.9	192,749	97.9	153,462	99.6
CA	554,710	7.1	192,513	49.4	1,987,290	88.3	1,198,000	97.4
FL	239,790	5.5	111,404	68.9	1,276,250	96.0	1,028,672	98.8
IA	41,667	2.6	11,333	28.3	144,532	69.1	140,035	80.9
MA	83,107	3.6	33,745	44.9	410,153	93.5	322,941	97.7
MS	32,490	12.5	20,579	83.0	199,393	92.4	137,348	91.6
NE	28,263	100.0	9,975	100.0	99,495	100.0	74,366	100.0
NM	28,409	95.6	11,862	98.8	104,959	97.6	61,479	97.5
NY	268,909	95.6	111,087	97.8	1,321,384	96.0	877,212	99.5
UT	56,613	57.9	16,291	43.4	143,668	92.7	64,199	97.6
VT	6,117	9.8	1,516	76.9	24,627	87.0	19,951	86.2
WA	88,854	100.0	25,729	100.0	325,198	100.0	208,288	100.0

State	State Emergency Department Databases, 2011							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	359,619	27.7	2,244,368	44.6	6,206,312	87.3	1,313,903	96.0
FL	217,616	44.8	1,439,944	66.1	4,355,028	95.2	873,179	97.8
IA	31,988	18.1	221,375	37.4	617,924	78.6	170,191	87.6
MA	49,307	20.2	445,314	45.4	1,709,624	89.2	304,156	92.8
NE	19,352	100.0	106,494	100.0	270,207	98.9	69,745	99.9
NY	196,430	86.5	1,371,705	88.2	4,379,886	95.9	697,750	98.8
UT	20,464	10.2	140,825	33.4	428,533	92.2	76,051	97.2
VT	3,841	33.7	41,229	64.0	158,535	80.7	39,074	81.3

State	State Ambulatory Surgery and Services Databases, 2011							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	8,610	26.9	179,433	46.8	1,209,851	91.9	675,698	96.5
FL	10,976	52.0	139,903	66.3	1,513,911	95.5	1,258,393	97.4
IA	3,584	13.6	33,036	30.4	260,810	63.8	147,611	71.9
NE	2,628	100.0	19,795	100.0	102,813	99.9	53,024	100.0
NY	7,893	98.9	132,406	98.6	1,376,624	97.7	655,202	99.8
UT	5,391	12.6	39,283	38.2	230,641	92.4	100,677	92.6
VT	8,489	31.5	87,299	68.7	918,190	83.5	563,044	83.4

State	State Inpatient Databases, 2010							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	42,862	60.9	20,059	86.4	194,909	98.3	154,126	99.7
CA	564,911	7.3	198,865	51.2	1,998,102	88.3	1,208,263	97.5
FL	241,766	6.1	110,016	70.7	1,266,088	96.0	1,022,113	98.9
IA	42,516	3.0	11,231	29.0	145,762	65.3	141,746	75.4
MA	82,779	3.8	33,784	47.9	408,199	94.0	321,250	97.9
MS	30,162	15.8	19,435	86.3	191,645	96.0	134,395	98.1
NC	130,423	9.1	46,206	50.7	569,558	73.6	383,129	76.4
NE	29,062	100.0	9,826	100.0	101,992	100.0	78,045	100.0
NM	29,017	96.1	12,462	98.7	105,757	97.9	61,821	97.7
NY	273,146	95.4	113,481	96.9	1,340,145	94.0	885,575	99.0
UT	57,626	57.2	15,288	42.0	140,726	92.7	60,770	97.4
WA	89,103	100.0	27,294	100.0	327,885	100.0	207,486	100.0

State	State Emergency Department Databases, 2010							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	366,288	28.4	2,188,741	45.6	5,968,456	87.4	1,213,911	96.0
FL	215,640	47.4	1,390,084	67.8	4,159,865	95.2	836,457	97.9
IA	31,946	16.4	210,199	37.3	597,782	75.7	167,297	83.2
MA	50,342	22.6	438,302	48.0	1,695,199	89.5	294,067	93.2
NC	92,944	30.9	670,615	51.0	2,458,847	76.7	401,007	78.5
NE	20,154	100.0	106,199	100.0	261,563	98.6	75,385	99.8
NY	196,812	86.7	1,326,578	88.3	4,149,237	94.2	663,130	98.3
UT	20,800	11.2	138,206	34.0	415,136	91.6	74,589	96.7

State	State Ambulatory Surgery and Services Databases, 2010							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	9,236	28.3	182,226	47.3	1,282,536	92.0	701,146	96.1
FL	11,716	51.4	143,722	68.1	1,574,434	96.2	1,284,377	98.1
IA	3,885	11.8	32,446	29.2	253,248	59.1	150,379	64.7
NC	13,034	22.7	177,358	46.1	973,012	64.2	475,476	61.1
NE	2,747	100.0	19,438	100.0	103,327	100.0	54,386	99.9
NY	5,580	98.4	85,706	92.1	852,743	82.7	412,752	98.8
UT	5,481	12.0	38,337	37.0	221,286	88.5	96,834	90.0

State	State Inpatient Databases, 2009							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	45,214	65.0	22,142	85.6	195,911	98.3	154,867	99.8
CA	585,245	5.5	206,366	52.1	2,000,666	88.2	1,192,387	97.6
FL	251,382	5.8	111,325	71.9	1,261,315	95.9	982,054	98.9
IA	43,684	2.8	12,373	30.5	148,533	65.9	144,365	77.1
NC	135,811	9.4	48,804	51.7	567,120	73.7	375,068	77.0
NE	30,303	99.9	10,174	100.0	98,770	100.0	76,907	100.0
NM	30,239	97.2	13,523	98.9	107,868	98.0	60,743	97.6
NY	279,861	95.3	120,927	96.9	1,363,573	93.6	897,257	98.8
UT	59,507	55.6	15,545	40.3	142,322	92.9	59,351	98.2
WA	92,138	100.0	27,809	100.0	328,727	100.0	203,993	100.0

State	State Emergency Department Databases, 2009							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	394,223	28.9	2,375,278	46.6	5,933,157	87.3	1,172,607	96.0
FL	235,897	51.4	1,480,495	70.8	4,040,055	95.2	776,721	98.0
NC	105,337	32.1	751,923	52.4	2,476,967	76.5	440,515	78.9
NE	22,781	100.0	116,277	100.0	255,074	99.1	74,264	99.8
NY	208,321	85.3	1,465,813	87.8	4,127,941	94.2	635,494	98.4
UT	24,565	9.6	158,767	33.1	433,969	91.8	72,088	96.9

State	State Ambulatory Surgery and Services Databases, 2009							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	9,843	28.8	184,331	48.6	1,416,182	91.6	765,289	96.0
FL	14,326	46.2	138,063	71.0	1,616,675	96.6	1,301,674	98.3
NC	13,669	26.2	175,108	47.4	950,338	64.4	451,080	61.5
NE	1,709	100.0	16,610	100.0	97,129	99.9	52,852	99.7
NY	8,775	99.0	141,282	91.7	1,409,196	82.8	666,165	99.3
UT	5,293	9.4	36,795	35.3	191,160	94.0	78,060	94.6

State	State Inpatient Databases, 2008							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	45,795	68.7	21,475	85.8	197,452	98.3	160,394	99.8
CA	612,708	5.5	199,522	54.3	2,006,150	87.1	1,199,350	97.5
FL	260,731	5.6	108,399	72.0	1,234,624	95.2	967,968	98.9
NC	140,827	9.6	47,067	52.7	569,233	71.4	376,886	74.8
NE	28,195	99.8	9,601	100.0	95,259	99.8	82,421	100.0
NY	278,026	86.4	116,578	88.7	1,343,000	91.6	891,549	98.7
UT	62,002	57.4	15,082	42.6	142,735	92.9	59,648	98.5
WA	94,611	100.0	27,679	100.0	325,906	100.0	204,116	100.0

State	State Emergency Department Databases, 2008							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	378,350	29.6	1,999,440	47.1	5,534,462	86.7	1,120,595	95.6
FL	219,843	49.2	1,252,403	69.7	3,763,524	94.9	742,369	98.0
NC	103,572	32.5	664,834	51.7	2,270,011	73.2	367,417	74.7
NE	20,622	100.0	100,447	99.9	237,883	99.1	71,622	99.7
NY	203,056	83.5	1,313,716	86.8	3,895,188	93.8	607,766	98.3
UT	25,313	10.0	148,214	35.3	440,314	91.2	73,309	96.6

State	State Ambulatory Surgery and Services Databases, 2008							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	10,405	32.3	191,981	49.2	1,712,134	90.1	937,650	94.4
FL	14,864	47.3	136,925	71.2	1,670,260	96.9	1,310,561	98.5
NC	14,416	30.3	162,534	48.5	905,289	65.5	425,546	66.7
NE	1,669	100.0	15,597	100.0	88,762	99.9	51,071	99.8
NY	9,104	90.2	138,337	84.7	1,324,454	82.0	629,602	99.0
UT	4,733	7.6	34,865	38.2	188,989	95.1	71,356	96.6

State	State Inpatient Databases, 2007							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	47,119	82.2	23,008	85.5	198,589	98.2	159,252	99.8
AZ	115,917	96.7	37,643	97.4	384,047	98.2	237,973	99.5
CA	630,527	5.0	196,930	55.6	2,002,684	87.1	1,182,402	97.5
FL	269,963	6.0	109,591	72.4	1,230,170	94.9	953,606	98.9
NC	141,105	13.5	46,588	54.9	562,470	72.6	370,135	76.0
NE	29,066	99.9	9,990	100.0	93,990	99.8	79,577	100.0
NV	43,547	36.5	11,155	41.0	153,146	81.7	81,679	88.1
NY	277,128	86.4	122,460	89.8	1,334,075	94.2	874,927	99.3
UT	61,268	58.7	15,481	45.5	142,132	93.4	60,051	98.7
WA	94,506	41.4	27,658	42.1	321,773	33.7	200,910	34.0

State	State Emergency Department Databases, 2007							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AZ	85,748	96.0	389,029	93.7	1,075,792	95.7	208,971	99.1
CA	377,042	29.5	1,962,937	47.3	5,376,070	85.8	1,075,002	93.5
FL	219,728	46.8	1,222,668	69.0	3,625,210	94.9	702,066	98.1
NC	103,220	34.4	660,039	52.1	2,171,606	72.1	349,324	72.0
NE	21,571	100.0	104,787	100.0	227,030	99.4	65,139	99.8
NY	197,217	84.4	1,267,357	87.4	3,558,764	96.1	559,385	99.1
UT	25,263	10.6	150,736	37.1	440,707	91.6	70,263	96.6

State	State Ambulatory Surgery and Services Databases, 2007							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	10,375	37.0	187,822	52.2	1,815,642	90.8	984,808	94.4
FL	14,927	43.7	138,427	73.4	1,683,974	97.1	1,257,278	98.8
NC	12,036	24.6	160,090	35.8	889,708	47.7	392,402	50.0
NE	2,188	100.0	17,612	100.0	98,395	99.8	61,982	99.8
NY	8,662	90.2	125,344	92.8	1,135,203	98.6	527,862	99.7
UT	4,728	7.8	33,770	40.6	188,478	95.8	69,614	97.1

State	State Inpatient Databases, 2006							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	46,640	80.6	22,714	84.9	198,642	98.1	161,634	99.7
AZ	115,758	96.9	37,377	97.3	378,767	98.2	236,600	99.5
CA	626,137	5.1	200,097	57.8	1,978,294	87.7	1,192,374	97.8
FL	268,967	6.5	111,248	74.2	1,212,492	95.0	958,311	98.9
NC	137,025	14.8	46,672	54.3	561,755	71.4	370,733	73.8
NE	29,042	99.9	9,820	100.0	93,619	99.4	80,125	99.9
NV	42,283	46.2	10,218	53.2	150,015	93.8	82,638	98.4
NY	286,449	86.1	125,410	89.5	1,358,888	94.2	888,247	99.5
UT	59,457	57.0	14,848	48.8	138,916	94.2	59,173	99.0
WA	92,143	100.0	27,567	100.0	315,643	100.0	198,675	100.0

State	State Emergency Department Databases, 2006							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AZ	80,522	96.2	368,327	94.1	1,017,179	96.1	201,996	99.1
CA	358,400	30.8	1,884,820	48.4	5,232,509	85.6	1,052,777	94.0
FL	221,362	51.9	1,275,674	71.5	3,649,681	94.8	703,314	98.1
NE	21,619	100.0	102,007	100.0	213,182	99.1	63,709	99.5
NY	203,079	83.8	1,286,823	87.1	3,566,588	96.0	563,590	99.0
UT	25,667	13.8	151,413	38.9	429,040	92.4	68,706	97.1

State	State Ambulatory Surgery and Services Databases, 2006							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	10,650	38.3	183,118	54.7	1,750,397	92.4	930,637	95.4
FL	15,261	43.7	135,629	75.8	1,616,070	97.4	1,178,266	99.0
NC	14,196	38.9	187,592	54.7	1,031,812	70.1	389,829	68.7
NE	2,602	100.0	19,180	100.0	103,981	99.8	68,596	99.9
NY	7,314	91.1	115,314	93.1	1,025,678	98.7	460,268	99.7
UT	4,067	13.2	34,958	42.5	191,037	96.4	74,365	96.6

State	State Inpatient Databases, 2005							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	44,961	81.3	22,416	84.4	197,648	98.2	163,993	99.7
AZ	108,992	91.2	37,825	96.7	362,706	97.9	236,015	99.5
CA	615,356	5.3	207,292	59.9	1,966,134	88.5	1,201,198	98.0
FL	260,668	7.3	113,752	76.4	1,189,067	95.4	971,121	99.0
NC	132,902	14.1	47,134	55.5	547,597	69.5	367,859	71.7
NE	26,230	99.9	9,485	100.0	84,127	99.2	72,393	99.7
NV	39,840	52.2	10,426	64.1	145,126	96.2	80,561	98.6
NY	283,111	86.5	126,851	89.5	1,346,698	94.2	889,709	99.4
UT	57,452	57.6	15,555	50.1	135,239	94.1	60,387	96.7
WA	88,036	100.0	28,073	100.0	304,791	100.0	196,153	100.0

State	State Emergency Department Databases, 2005							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AZ	85,391	91.4	391,569	90.2	1,006,070	96.3	205,364	99.2
CA	355,358	32.9	1,946,166	50.5	5,219,670	84.8	1,038,997	93.7
FL	219,289	58.4	1,258,138	74.2	3,572,975	95.2	707,533	98.1
NE	18,896	100.0	92,697	100.0	185,203	99.2	55,838	99.6
UT	26,284	13.4	159,082	38.8	416,383	91.8	66,835	95.2

State	State Ambulatory Surgery and Services Databases, 2005							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
CA	10,768	41.5	179,092	57.0	1,693,309	93.6	907,690	96.2
FL	14,556	47.4	131,253	79.3	1,519,009	97.9	1,135,950	99.1
NC	13,136	42.7	184,258	54.7	1,009,371	67.1	369,422	66.1
NE	1,797	100.0	18,465	99.9	97,100	99.0	58,857	99.8
NY	13,576	84.9	146,023	90.4	1,092,680	98.3	461,604	99.6
UT	4,101	20.1	36,688	46.1	190,790	96.8	76,870	96.5

State	State Inpatient Databases, 2004							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AR	44,366	83.3	22,226	84.7	198,100	98.4	164,958	99.6
AZ	105,002	63.9	35,588	79.1	350,360	95.9	226,044	99.1
CA	609,724	5.7	208,575	63.8	1,956,858	89.8	1,182,267	98.3
FL	250,408	7.7	115,312	78.7	1,155,435	96.2	969,521	99.1
NC	129,195	14.4	46,549	53.2	541,367	67.5	365,218	69.9
NE	24,715	99.9	8,971	99.6	79,542	99.6	68,349	99.3
NV	37,137	55.7	9,798	69.1	134,496	96.9	76,500	98.8
NY	287,629	94.1	129,782	97.2	1,355,708	95.5	887,860	99.4
UT	57,090	58.4	14,877	47.0	134,435	94.4	59,778	96.5
WA	86,437	100.0	26,374	100.0	296,444	100.0	187,384	100.0

State	State Emergency Department Databases, 2004							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
NE	17,540	100.0	85,828	100.0	175,521	99.0	53,701	99.2
UT	25,253	15.1	147,329	39.1	394,328	92.1	61,390	95.6

State	State Ambulatory Surgery and Services Databases, 2004							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
FL	11,631	62.2	135,055	81.5	1,483,788	98.4	1,116,948	99.4
NC	10,419	42.4	155,996	52.6	883,811	64.3	348,958	63.8
NE	1,626	100.0	20,022	99.2	95,810	99.6	57,830	99.6
NY	6,750	98.2	103,963	98.2	907,074	98.8	420,631	99.4
UT	3,974	20.9	38,277	47.4	187,803	97.6	73,050	97.2

State	State Inpatient Databases, 2003							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
AZ	100,663	59.8	34,574	70.5	323,865	95.5	204,318	99.2
CA	609,241	6.7	215,088	66.1	1,951,540	90.2	1,203,490	98.4
NC	128,121	13.4	49,656	48.6	531,725	63.7	359,191	66.9
NE	27,920	100.0	10,270	99.8	91,634	99.7	76,520	100.0
NV	35,481	54.1	9,465	73.1	124,658	97.1	71,157	99.1
NY	279,299	94.9	124,294	97.1	1,271,074	96.1	843,251	99.8
UT	56,112	59.6	15,315	49.0	131,432	95.4	58,039	97.4
WA	85,069	100.0	26,824	100.0	291,333	100.0	185,787	100.0

State	State Emergency Department Databases, 2003							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
NE	21,333	100.0	99,660	100.0	196,442	98.7	57,874	99.9
UT	27,103	15.6	161,138	39.6	390,579	93.7	60,347	97.3

State	State Ambulatory Surgery and Services Databases, 2003							
	Age in Years							
	0		1-17		18-64		65+	
	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified	Number of Total Records	Percent Verified
NC	10,385	40.8	136,476	50.8	786,155	62.7	327,328	62.7
NE	2,171	100.0	27,813	99.6	115,113	99.3	65,771	99.5
NY	11,511	96.8	125,193	97.9	811,543	99.2	361,368	99.8
UT	3,810	24.1	35,423	47.8	173,395	97.7	67,226	96.8

APPENDIX E: SAS CODE FOR USAGE EXAMPLES

Usage Example #1: Assigning Patient Characteristics

```
** Assigning Attributes

** combined event data -- combine all available data types for the years;
** for simplicity, the example uses only 2 attributes: AGE and ZIPINC_QRTL;
data NE_CombinedEvents1;
  keep KEY visitLink daysToEvent AGE ZIPINC_QRTL
       events   sidEvents   seddEvents   sidSeddEvents   sasdEvents
       verified sidVerified seddVerified sidSeddVerified sasdVerified;
  merge NE_2005_DaysToEvent (in=_inDaysToEvent)
        NE_2006_DaysToEvent (in=_inDaysToEvent)
        NE_SID_2005_Core (in=_inSID_2005)
        NE_SID_2006_Core (in=_inSID_2006)
        NE_SASD_2005_Core (in=_inSASD_2005)
        NE_SASD_2006_Core (in=_inSASD_2006)
        NE_SEDD_2005_Core (in=_inSEDD_2005)
        NE_SEDD_2006_Core (in=_inSEDD_2006)
        end=lastObs;
  by KEY;

  ** indicator for all events;
  events = 1;
  if _inDaysToEvent then verified = 1;
  else                   verified = 0;

  ** indicator for SID events;
  if _inSID then do;
    sidEvents = 1;
    if _inDaysToEvent then sidVerified = 1;
    else                   sidVerified = 0;
  end;

  ** indicator for SEDD events;
  if _inSEDD then do;
    seddEvents = 1;
    if _inDaysToEvent then seddVerified = 1;
    else                   seddVerified = 0;
  end;

  ** indicator for SID/SEDD events;
  sidSeddEvents = max(sidEvents, seddEvents);
  sidSeddVerified = max(sidVerified, seddVerified);

  ** indicator for SASD events;
  if _inSASD then do;
    sasdEvents = 1;
    if _inDaysToEvent then sasdVerified = 1;
    else                   sasdVerified = 0;
  end;
run;


** sort the combined events into link order;
proc sort data=NE_CombinedEvents1;
  by visitLink daysToEvent;
run;
```

```
** use the first non-missing value for link ID attributes;
data NE_LinkAttribs2;
  keep visitLink foundAGE foundZIP;
  set NE_CombinedEvents1;
  by visitLink daysToEvent;
  where visitLink;  ** only process events with a visitLink ID;

  ** set attributes to missing for each new verified patient;
  length foundZIP $5;
  retain foundAGE foundZIP;
  if first.visitLink then do;
    foundAGE = .;
    foundZIP = '';
  end;

  ** select the first non-missing attribute;
  if foundAGE le .Z and AGE gt .Z then foundAGE = AGE;
  if foundZIP eq '' and ZIPINC_QRTL ne '' then foundZIP = ZIPINC_QRTL;

  ** create one set of attributes for each verified patient;
  if last.visitLink then output;
run;


** apply the consistent attributes to the event data;
data NE_CombinedEvents2;
  keep KEY visitLink daysToEvent AGE ZIPINC_QRTL
       events   sidEvents   seddEvents   sidSeddEvents   sasdEvents
       verified sidVerified seddVerified sidSeddVerified sasdVerified;
  merge NE_CombinedEvents1
        NE_LinkAttribs2 (in=_inLinkAttribs)
  by visitLink;

  ** apply uniform attributes to verified patient events;
  if _inLinkAttribs then do;
    AGE = foundAGE;
    ZIPINC_QRTL = foundZIP;
  end;
run;
```

Usage Example #2: Revisits for Selected Patients

```
** Example 2, Follow-up Care
** Program code

** diagnosis CCS code to examine and clean period (months);
%let condX = 50;        ** diabetes mellitus with complications;
%let cleanPeriod = 6;   ** number of months for "clean" period with no condX claims;
%let firstYear = 2005;  ** first year of data;
%let lastYear  = 2006;  ** last year of data;


** (1) combine event  and revisit data -- keep only events with the specified condition;
data NE_EventsCondX1;
  keep KEY visitLink daysToEvent servSetting YEAR DQTR AMONTH LOS;
  merge NE_2005_DaysToEvent (in=_inDaysToEvent)
        NE_2006_DaysToEvent (in=_inDaysToEvent)
        NE_SID_2005_Core (in=_inSID)
        NE_SID_2006_Core (in=_inSID)
        NE_SASD_2005_Core (in=_inSASD)
        NE_SASD_2006_Core (in=_inSASD)
        NE_SEDD_2005_Core (in=_inSEDD)
        NE_SEDD_2006_Core (in=_inSEDD)
        end=lastObs;
  by KEY;

  ** (1.a) limit data to linkable patients;
  if _inDaysToEvent and (_inSID or _inSASD or _inSEDD);

  ** service type indicator;
  select;
    when (_inSID)  servSetting = 'IP';
    when (_inSEDD) servSetting = 'ED';
    when (_inSASD) servSetting = 'AS';
    otherwise;
  end;

  ** (1.b) grab all events with condX;
  if _n_ eq 1 then put "Searched CCS diagnoses codes for values of '&condX'";
  array DXCCS {*} DXCCS1-DXCCS15;
  do i = 1 to NDX;
    if DXCCS{i} eq &condX then output NE_EventsCondX1;
  end;
run;


** (2) sort the combined events into visitLink (patient), service sequence order;
proc sort data=NE_EventsCondX1;
  by visitLink daysToEvent;
run;
```

```
** (3) find people with 2+ condX related events,
** calculate days between the 1st and 2nd event;
data NE_EventsCondX2;
  keep visitLink patientCount patientEvents servSetting1 servSetting2 days_E1toE2;
  set NE_EventsCondX1
      end=lastObs;
  by visitLink daysToEvent;

  retain patientCount 1;
  label patientCount  = 'count of patients'
        patientEvents = 'number of events for patient'
        servSetting1  = 'type of service for the initial event'
        servSetting2  = 'type of service for the second event'
        days_E1toE2   = 'days between initial and second event';

  retain _condX_events;
  _condX_events + 1;
  _dayLag = lag(daysToEvent);

  retain _patients _pat_w2Plus _pat_wCleanPeriod _cleanPeriod days_E1toE2 patientEvents 0
         servSetting1 servSetting2;
  if first.visitLink then do;
    _patients + 1;  ** count patients with condX;

    ** reset patient indicators and counters;
    days_E1toE2 = .;
    patientEvents = .;
    servSetting1 = '  ';
    servSetting2 = '  ';

    if not (first.visitLink and last.visitLink) then _pat_w2Plus + 1;

    ** (3.a.i) first claim must be after "cleanPeriod" month of first data year;
    if (YEAR gt &firstYear) or  /* assumes clean period (months) < 12 */
       ( YEAR eq &firstYear and
         DQTR gt &cleanPeriod/3 and
         AMONTH gt &cleanPeriod and
         LOS lt &cleanPeriod*30 ) then do;
      _cleanPeriod = 1;  ** indicator that patient had clean period;
      servSetting1 = servSetting;
      patientEvents = 1;  ** counter for number of patient events;
    end;  /* end-if (clean period) */
    else _cleanPeriod = 0;
  end;  /* end-if (first visitLink) */

  else do;  /* not first visitLink */
    ** (3.b) revisit - clean period already found;
    if _cleanPeriod then do;
      patientEvents + 1;  ** count number of events for this patient;
      if patientEvents eq 2 then do;
        days_E1toE2 = daysToEvent - _dayLag;
        servSetting2 = servSetting;
        _pat_wCleanPeriod + 1;  ** count patients with a clean period;
      end;  /* end-if (second event) */
    end;  /* end-if (clean period) */

    else do;  /* clean not (yet) period found */
      ** (3.a.ii) no clean pd yet found - check lag days f/ clean pd between events;
      if _dayLag/30 gt &cleanPeriod then do;  /* 99% correct */
        _cleanPeriod = 1;  ** indicator that patient had clean period;
        servSetting1 = servSetting;
        patientEvents = 1;  ** counter for number of patient events;
      end;  ** end-if (lag clean period);
    end;  /* end-else (clean period not found) */
  end;  /* end-else (not first visitLink) */

  ** output one obs per visitLink;
  if last.visitLink and patientEvents ge 2 then output;
```

```
  ** (3.c) summarize processing;
  if lastObs then do;
    put '================================================================================';
    put "Processing summary -- events with diagnosis (CCS) category: &condX";
    put '-started with ' _condX_events 'events (total)';
    put '          for ' _patients 'patients.';
    put '-there were ' _pat_w2Plus "patients with 2+ DXCCS '&condX' events";
    put "-clean period, no claim with DXCCS '&condX' for at least &cleanPeriod months";
    put ' before the first claim: ' _pat_wCleanPeriod ;
    put '================================================================================';
    put /;
  end;
run;


proc format;
  picture pctfmt low-high='009 %';
run;


** (4) statistics for the number of condX related events;
title1 "Number of DXCCS &condX Events - Distribution";
title2 "for Patients with Multiple DXCCS &condX Events";
proc means data=NE_EventsCondX2 maxdec=2 mean p25 p50 p75 max;
  var patientEvents;
run;


** (5) summarize revisits by the initial and second service settings;
title "Count and Days for DXCCS '&condX' Patients with Multiple Hospital Events";
proc tabulate data=NE_EventsCondX2 format=comma12.;
  class servSetting1 servSetting2 /descending;
  var patientCount;
  table (servSetting1 all),
        (servSetting2 all)*
        (patientCount*sum*f=comma12.
         days_E1toE2*(mean median)*f=8.1);
run;
```

Output

Number of DXCCS 50 Events - Distribution
for Patients with Multiple DXCCS 50 Events

The MEANS Procedure

Analysis Variable : patientEvents number of events for patient

Mean	25th Pctl	50th Pctl	75th Pctl	Maximum
4.14	2.00	3.00	4.00	97.00

Count and Days for DXCCS '50' Patients with Multiple Hospital Events

	type of service for the second event											
	IP			ED			AS			All		
	count of patients	days between initial and second event		count of patients	days between initial and second event		count of patients	days between initial and second event		count of patients	days between initial and second event	
	Sum	Mean	Median	Sum	Mean	Median	Sum	Mean	Median	Sum	Mean	Median
initial event												
IP	1,850	61.6	6.0	367	102.8	35.0	157	140.3	111.0	2,374	73.2	14.0
ED	254	122.0	75.0	299	93.3	34.0	40	123.0	86.5	593	107.6	52.0
AS	158	151.3	120.0	24	114.0	115.5	466	41.6	14.0	648	71.0	21.0
All	2,262	74.7	16.0	690	99.1	37.0	663	69.9	21.0	3,615	78.5	21.0

Usage Example #3: Preceding Visits to any Hospital Setting for Selected Patients

```
** Example 3, Preceding Events
** Program code

** procedure CCS code to examine;
%let servX = 44;        ** CABG;


** combine event  and revisit data -- flag events with the specified service;
data TN_AllEvents;
  keep KEY visitLink daysToEvent servX dxCcs1 servSetting events;
  merge TN_2005_DaysToEvent (in=_inDaysToEvent)
        TN_2006_DaysToEvent (in=_inDaysToEvent)
        TN_SID_2005_Core (in=_inSID)
        TN_SID_2006_Core (in=_inSID)
        TN_SASD_2005_Core (in=_inSASD)
        TN_SASD_2006_Core (in=_inSASD)
        TN_SEDD_2005_Core (in=_inSEDD)
        TN_SEDD_2006_Core (in=_inSEDD)
        end=lastObs;
  by KEY;
  if _inSID  or _inSASD or _inSEDD;

  ** flag events with the specified service;
  array prCcs {*} prCcs1-prCcs30;
  do i = 1 to dim(prCcs);
    if prCcs{i} eq &servX then do;
      servX = 1;
      leave;
    end;
  end;

  array cptCcs {*} cptCcs1-cptCcs30;
  do i = 1 to dim(cptCcs);
    if cptCcs{i} eq &servX then do;
      servX = 1;
      leave;
    end;
  end;

  ** service type indicator;
  select;
    when (_inSID)  servSetting = 'IP';
    when (_inSEDD) servSetting = 'ED';
    when (_inSASD) servSetting = 'AS';
    otherwise;
  end;

  ** counters;
  retain events 1;
  if _inSID then sidEvents = 1;
  if _inSEDD then seddEvents = 1;
  if _inSASD then sasdEvents = 1;
run;


** sort the combined events into visitLink (patient), service sequence order;
proc sort data=TN_AllEvents out=TN_OrderedEvents (index=(servX));
  by visitLink daysToEvent;
run;
```

```
** find people who received the topic service;
proc sql;
  create table TN_TopicPop as
    select visitLink,
           min(daysToEvent) as firstService
    from TN_OrderedEvents
    where servX eq 1
    group by visitLink;
quit;


** for people with the service, select all events prior to the service;
data TN_PriorEvents;
  merge TN_OrderedEvents (in=_inEventData)
        TN_TopicPop (in=_inTopicPop);
  by visitLink;
  if _inTopicPop;
run;


** summarize prior events by primary diagnosis and setting;
title 'Count of patients by Primary Diagnosis and Service Setting';
proc tabulate data=TN_PriorEvents format=comma12.;
  class dxCcs1 servSetting;
  format DXCCS1 FDCCSPDX.;
  var events;
  table (all DXCCS1),
        (all servSetting)*(events*sum);
run;
```

Output

Count of patients by Primary Diagnosis and Service Setting

		servSetting		
	All	AS	ED	IP
	events	events	events	events
	Sum	Sum	Sum	Sum
All	613,541	87,348	292,336	233,857
CCS: principal diagnosis				
1: Infectious and Parasitic DX	8,351	273	6,726	1,352
2: Neoplasms	27,723	24,110	221	3,392
3: Endocr, Nutri, Metab, Immun DX	7,074	1,626	3,095	2,353
4: Dx of Blood, Blood-Forming Organs	2,689	416	1,510	763
5: Mental Disorders	8,004	216	5,939	1,849
6: Dx of Nervous System, Sense Organs	46,362	9,754	34,891	1,717
7: Dx of Circulatory System	72,483	13,626	18,290	40,567
8: Dx of Respiratory System	59,376	7,848	43,290	8,238
9: Dx of Digestive System	30,309	5,802	18,551	5,956
10: Dx of Genitourinary System	17,253	3,260	11,766	2,227
11: Complic Preg, Birth, Puerperium	15,170	2,244	4,302	8,624
12: Dx of Skin and Subcutaneous Tissue	8,553	826	6,626	1,101
13: Dx of Musculoskel, Connective Tissue	25,937	6,453	16,583	2,901
14: Congenital Anomalies	2,098	917	180	1,001
15: Perinatal Conditions	145,485	810	3,227	141,448
16: Injury and Poisoning	90,679	3,639	79,694	7,346
17: Other Conditions	45,995	5,528	37,445	3,022